# Solid-phase glycopeptide synthesis of tyrosine-glycosylated glycogenin fragments as substrates for glucosylation by glycogenin

**Anita M. Jansson,[a] Knud J. Jensen,[a] Morten Meldal,*[,a] Joseph Lomako,[b] Wieslawa M. Lomako,[b] Carl Erik Olsen[c] and Klaus Bock[a]**

[a] *Department of Chemistry, Carlsberg Laboratory, Gamle Carlsberg Vej 10, DK-2500 Valby, Copenhagen, Denmark*
[b] *Department of Biochemistry and Molecular Biology, University of Miami School of Medicine, PO Box 016129, Miami, FL 33101-6129, USA*
[c] *Department of Chemistry, Royal Veterinary and Agricultural University of Denmark, Thorvaldsensvej 40, DK-1871 Frederiksberg C, Denmark*

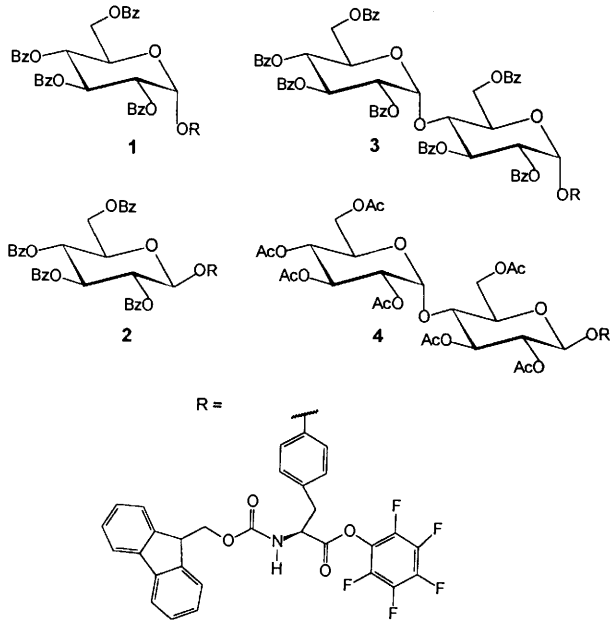
**The four building blocks $N^{\alpha}$-Fmoc-Tyr($Bz_4$-α-D-Glc)-*O*Pfp 1, $N^{\alpha}$-Fmoc-Tyr($Bz_4$-β-D-Glc)-*O*Pfp 2, $N^{\alpha}$-Fmoc-Tyr[$Bz_4$-α-D-Glc-(1→4)-$Bz_3$-α-D-Glc]-*O*Pfp 3 and $N^{\alpha}$-Fmoc-Tyr[$Ac_4$-α-D-Glc-(1→4)-$Ac_3$-β-D-Glc]-*O*Pfp 4 were used in continuous-flow solid-phase synthesis of four 25-amino-acid-long glycopeptides 5–8 related to glycogenin. The unglycosylated peptide 9 was synthesized as well, and the five glycogenin fragments were tested as acceptors for glucosyl transfer by recombinant glycogenin.**

## Introduction

Most proteins found in nature are glycosylated,[1,2] and the glycan part of a glycoprotein can be either N-linked to asparagine, O-linked to serine, threonine, 5-hydroxylysine, 4-hydroxyproline or tyrosine,[2–4] or C-linked to tryptophan.[5] The isolation of glycoproteins and glycopeptides from natural sources is generally difficult and the isolated glycoconjugates often show a microheterogeneity in the glycan moiety.[1] Therefore, access to efficient methodologies for chemical synthesis of well defined glycopeptides is of importance.

O-Glycosylated tyrosine, as well as glycopeptides with the glycan part linked to the phenolic hydroxy group of tyrosine, have been found at various sites in nature.[6–13] In glycogenin, the 332-amino-acid-long primer protein for the biosynthesis of glycogen, the aromatic hydroxy group of Tyr-194 is glucosylated.[7,14] Once the first glucose is attached, an average of up to seven further glucose units are added. All eight glucose units are proposed to be added by self-glucosylation.[15] The elongation of this primer is then catalysed by glycogen synthase, and it is likely that this elongation is facilitated by the interaction between glycogenin and glycogen synthase, which exist as a 1:1 complex.[16] The anomeric configuration of the glucosyl residue linked to Tyr-194 in glycogenin was proposed to be α on the basis of the splitting of the glycogen–glycogenin bond by isoamylase.[17] To confirm the configuration of the glycosidic linkage to tyrosine in glycogenin, it was desirable to have access to well defined fragments of glycogenin in the vicinity of the glycosylation site with glucose and maltose α- and β-linked to tyrosine, as well as the corresponding unglycosylated peptide. Currently the only complete recorded synthesis of tyrosine-glycosylated peptides describes the preparation of di-, tetra- and penta-peptides with a glucosylated tyrosine at the N-terminus.[18,19] The peptides were prepared either by direct glycosylation of peptides, or by the glycosylation of $N^{\alpha}$-Z-Tyr-*O*Bu$^{t}$,† followed by deprotection of the carboxy function and incorporation of the glycosylated tyrosine derivative into a peptide in the last step of a solution-phase synthesis.

The fluoren-9-ylmethoxycarbonyl/pentafluorophenyl (Fmoc/Pfp) ester building-block strategy for synthesis of N- and O-linked glycopeptides has been found to be most versatile.[20] It involves direct glycosylation of an $N^{\alpha}$-Fmoc amino acid Pfp ester, and the use of this glycosylated building block in solid-phase peptide synthesis (SPPS).[20] This strategy has now been used for the assembly of four tyrosine-glycosylated glycogenin fragments **5–8**. Synthesis of the four building blocks $N^{\alpha}$-Fmoc-Tyr($Bz_4$-α-D-Glc)-*O*Pfp **1**, $N^{\alpha}$-Fmoc-Tyr($Bz_4$-β-D-Glc)-*O*Pfp **2**, $N^{\alpha}$-Fmoc-Tyr[$Bz_4$-α-D-Glc-(1→4)-$Bz_3$-α-D-Glc]-*O*Pfp **3** and $N^{\alpha}$-Fmoc-Tyr[$Ac_4$-α-D-Glc-(1→4)-$Ac_3$-β-D-Glc]-*O*Pfp **4** has been reported previously,[21] as well as synthesis of other glycosyltyrosine building blocks for use in SPPS.[22]




## Results and discussion

The secondary structure of unglycosylated glycogenin (apo-glycogenin) in the vicinity of the glycosylation site at Tyr-194 was predicted with a computer program based on the 3-dimensional GOR method.[23] According to this prediction a β-strand from Phe-184 to Tyr-194 and an α-helix from Ala-199 to

† Z = benzyloxycarbonyl.

**Table 1** Amino acid analysis of compounds **5–9** (theoretical values in parenthesis)

| Amino acid | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|
| Ala (4) | 4.2 | 4.4 | 4.0 | 4.3 | 4.3 |
| Asp[a] (2) | 2.1 | 2.1 | 2.0 | 2.0 | 2.0 |
| Gly (1) | 1.0 | 1.1 | 1.1 | 1.1 | 1.1 |
| Ile (3) | 2.7 | 2.7 | 2.8 | 2.9 | 2.7 |
| Leu (2) | 2.2 | 2.1 | 2.1 | 2.1 | 2.1 |
| Lys (2) | 2.2 | 2.1 | 2.0 | 2.0 | 2.1 |
| Phe (3) | 3.2 | 3.1 | 3.0 | 3.1 | 3.1 |
| Pro (1) | 1.0 | 1.0 | 1.0 | 1.0 | 1.1 |
| Ser (4) | 3.7 | 3.6 | 4.2 | 3.9 | 3.8 |
| Tyr (3) | 2.7 | 2.7 | 2.9 | 2.7 | 2.8 |

[a] Asn was determined as Asp.

**Table 2** $^1$H NMR data of the anomeric protons in compounds **5–8**[a]

| Compound | Chemical shift ($\delta_H$) 1-H; 1′-H | $J_{1,2}$; $J_{1',2'}$ (Hz) |
|---|---|---|
| **5** (α-gluco) | 5.51 | 3.5 |
| **6** (β-gluco) | 4.96 | 7.5 |
| **7** (α-malto) | 5.54; 5.34 | 3.5; 3.6 |
| **8** (β-malto) | 4.99; 5.36 | 7.5; 3.5 |

[a] In $D_2O$–$CD_3CO_2D$ (1:1), ~1 mmol dm$^{-3}$.

Lys-208 were connected by a coil segment including Pro-198, indicating a probability of a β-turn preceding the glycosylation site. A sequence of 25 amino acids from Phe-184 to Lys-208 was selected as a synthetic target and the sequence is presented below. Four glycopeptides, as well as the unglycosylated analogue, were synthesized by SPPS employing the building blocks **1–4**.

OR
|
Ac-Phe$^{184}$-Ile-Tyr-Asn-Leu-Ser-Ser-Ile-Ser-Ile-Tyr$^{194}$-Ser-Tyr-Leu-Pro-Ala-Phe-Lys-Ala-Phe-Gly-Ala-Asn-Ala-Lys$^{208}$-$NH_2$

**5** R = α-D-glucopyranosyl
**6** R = β-D-glucopyranosyl
**7** R = α-D-glucopyranosyl-(1→4)-α-D-glucopyranosyl
**8** R = α-D-glucopyranosyl-(1→4)-β-D-glucopyranosyl
**9** R = H

Sequence of the 25-amino-acid-long glycogenin fragments glycosylated at Tyr-194

The four glyco-analogues of glycogenin, compounds **5–8**, were synthesized by standard Fmoc SPPS technique, on a custom-made continuous-flow synthesizer equipped with a solid-phase photometer, with dimethylformamide (DMF) as solvent. PEGA-Resin[24,25] [poly(ethyleneglycol)–poly(dimethylacrylamide) copolymer] was derivatized with 4-[α-(fluoren-9-ylmethoxycarbonylamino)-2,4-dimethoxybenzyl]phenoxyacetic acid[26] (Rink-linker) by the *O*-(1*H*-benzotriazol-1-yl)-*N*,*N*,*N*′,*N*′-tetramethyluronium tetrafluoroborate (TBTU) procedure.[27] This linker gives the C-terminal amide by acidic cleavage of the peptide from the resin. The side-chains of Ser, Thr, and unglycosylated Tyr were protected as their *tert*-butyl ethers, while Lys was $N^{\varepsilon}$-*tert*-butoxycarbonyl (Boc)-protected and the side-chain of Asn was N-tritylated (Trt). In the syntheses of the α-,β-gluco and the β-malto glycopeptides (**5**, **6** and **8**) the amino acids were coupled as 4-oxo-3,4-dihydro-1,2,3-benzotriazin-3-yl- (Dhbt) esters (2.5 equiv.), with the exceptions of $N^{\alpha}$-Fmoc-Asn(Trt)-OH, which was incorporated by the TBTU procedure, and the glycosylated building blocks **1**, **2** (2.5 equiv.) and **4** (2 equiv.), which were coupled with 3-hydroxy-4-oxo-3,4-dihydro-1,2,3-benzotriazine (Dhbt-OH) catalysis. In the synthesis of the α-maltose analogue **7**, all $N^{\alpha}$-Fmoc amino acids were coupled as Pfp esters[28] (3 equiv.) and only one equivalent of the building block **3** was used. Therefore, residual amino groups were capped with acetic anhydride after the coupling of compound **3**. [The truncated glycopeptide Ac-glycogenin(193–208)-$NH_2$ was also detected by electrospray mass spectrometry (ESMS) in the final crude product and could be isolated in ~3% yield.] Dhbt-OH (1 equiv.) was added as an auxiliary nucleophile when coupling steps were performed with Pfp esters. The acylation steps were monitored at 440 nm to record the coloration of the resin due to formation of an ion-pair between Dhbt-OH and free α-amino groups. This allows feedback control of the acylation reactions.[29,30] However, real-time monitoring with Dhbt-OH was found to be problematic due to some persistent coloration of the PEGA-resin, and therefore fixed coupling times were used. In early experiments on the synthesis of glycogenin-(184–208) fragment on Pepsyn KA resin, (with free C- and N-terminals) real-time monitoring of the acylation reactions revealed that the first 10 amino acids couple easily, while the following nine acids require prolonged reaction time, and the remaining six acids react very slowly.[31] Acylation reactions on PEGA-resin are in general faster than on Pepsyn resin, but as a precaution the acylation times on PEGA resin for all four glycopeptides were set to 90 min for the first ten amino acids, and 180 min for the next eight amino acids, with the exception of the glycosylated building blocks which were coupled overnight. The last six amino acids were allowed to react for 360 min for complete incorporation. Fmoc-deprotection was performed with 20% piperidine in DMF, and was monitored with a UV-detector by measuring the absorbency of the effluent at 312 nm during treatment with base. The time for Fmoc-deprotection of the first 14 amino acids was set to 10 min, for the following nine amino acids to 15 min, and for the last two amino acids to 30 min. The reduced rates of acylation and Fmoc-deprotection might be due to aggregation of the protected peptides on the resin. This is in accord with the prediction of a β-strand from Phe-184 to Ser-192, since aggregation phenomena on the resin, have in many cases, been ascribed to formation of a β-sheet.[32] However, the reduced rate of acylation and Fmoc-deprotection during synthesis was not as significant on the PEGA resin as on the Pepsyn resin.

After removal of the final $N^{\alpha}$-Fmoc group, the N-terminus was acetylated with Ac-*O*-Dhbt[33] (**5**, **6** and **8**) or acetic anhydride (**7**). The resin was washed with dichloromethane and dried before the peptide was cleaved from the resin with 95% aq. trifluoroacetic acid (TFA), followed by the addition of acetic acid and concentration. The crude glycopeptides were subjected to the final methoxide transesterification deprotection without further purification. Since the protected products were not fully soluble in methanol, the deprotection with methoxide was followed by analytical HPLC, and the reaction time was prolonged to 14–48 h. The products were purified by preparative reversed-phase HPLC and all purified products were analysed by analytical HPLC, amino acid analysis (Table 1) and ESMS. The anomeric configuration of the carbohydrate residues was confirmed by $^1$H NMR spectroscopy in $D_2O$–$CD_3CO_2D$ (1:1). Chemical shifts and coupling constants are presented in Table 2 and they all have the expected values.[34] Owing to low solubility and losses during the purification process, the overall yields of the purified glycopeptides **5**, **6** and **8** were only 4, 9 and 6%, respectively, while the overall yield of the α-maltosylated peptide **7** could be increased to 33% by dissolution of the crude sample in a small volume of aq. TFA before purification by HPLC. The unglycosylated peptide **9** was synthesized on a Synergy Personal Peptide Synthesizer in 11% overall yield. The yields are based on the loading of the resin as measured by a spectrophotometric method.[25]

Although only one equivalent of the glycosylated building block **3** was used in the synthesis of the α-maltosylated peptide, the yield of the final deprotected product **7** was better than those for products **5**, **6** and **8**. This is probably in part due

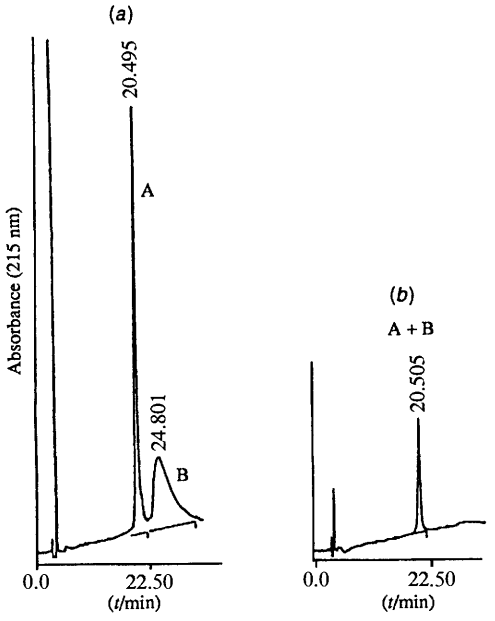



**Fig. 1** Analytical HPLC chromatograms of (*a*) a purified sample of β-maltosylated fragment **8** (20 μg injected), (*b*) a 1:1 mixture of samples from peak A and from peak B after preparative separation of A and B (reversed phase; linear gradient of 30–60% B for 30 min)

to the use of Pfp esters, but mostly to the modified purification procedure, where the crude deprotected product was dissolved in a small volume of 50% aq. TFA rather than a larger volume of aq. acetic acid. In addition to the problems due to low solubility, the glycopeptides also had a tendency to aggregate on the HPLC column, giving two or more peaks from a purified sample. An illustration of the phenomenon is shown in Fig. 1 where an analytical HPLC chromatogram of purified β-maltosylated peptide **8** is presented. Collection of fractions of the two peaks A and B was followed by reinjection of a 1:1 sample of A and B. This gave only one peak, with the same retention time as A. Peak A and peak B also gave identical spectra with ESMS. These results indicate that the two peaks in the HPLC chromatogram are due to different aggregates of the same compound. The tendency of the glycopeptides to aggregate also gave problems in the preparation of NMR samples. The best result was achieved by dissolution of an aliquot (2 mg) of a glycopeptide in a 1:1 mixture of $D_2O$–$CD_3CO_2D$ (~ 1 mmol $dm^{-3}$).

In an early attempt to improve the solubility, the crude protected products, in this case with free C- and N-terminals, were dissolved in a mixture of methanol and freshly distilled DMF before transesterification with methoxide.[31] However, this procedure resulted in partial mono- and di-formylation of the glycopeptides, and therefore DMF should not be used as a co-solvent in the deacetylation step. The formylated final crude products were characterized by liquid secondary-ion mass spectrometry (LSIMS) where peaks were detected at *m/z* values 28 and 56 higher than expected. Analysis of the sequencing data obtained from aliquots of deacetylated glycopeptides revealed that, while no unusual phenyl thiohydantion (PTH)-amino acid occurred when sequencing glycopeptides which still carried the protecting groups on the glycan, an unusual PTH-amino acid with a considerably shorter retention time was released together with the expected PTH-Lys. On the basis of these results it seems inevitable to conclude that a partial formylation of the $N^{\varepsilon}$-amino group of the Lys residues (possibly also of the terminal $N^{\alpha}$-amino group) occurs during the final deprotection if DMF is used as a co-solvent for the methoxide-catalysed transesterification reaction.

The four glycopeptides **5**–**8**, as well as the unglycosylated peptide **9**, were tested with recombinant glycogenin[15] in a transglucosylation reaction with labelled uridine 5′-diphospho-(UDP)-glucose. The results are shown in Table 3 and are expressed as percentage of the substrate undergoing glucosylation. Although the extent to which the substrates reacted is very low (< 0.05%), the maltosides are better acceptors than the glucosides. For comparison, *p*-nitrophenyl α-glucoside and *p*-nitrophenyl α-maltoside were tested as acceptors,[35] and were found to react even more slowly at a 10-times higher concentration. Recently, dodecyl β-maltoside was reported as being 30-times more active as a glycogenin substrate than was *p*-nitrophenyl α-maltoside.[36] By contrast, the unglucosylated peptide **9** did not act as a substrate, a result that others have also reported when using glycogenin(182–200)[16] and glycogenin(178–207)[37] as substrates. While this result was interpreted[16,37] as possibly implying that the addition of the first glucose residue to Tyr-194 is not autocatalytic, Alonso *et al.*[15] have demonstrated that recombinant glucose-free glycogenin does undergo autoglucosylation, adding an average of eight glucose residues, and that the addition of the first glucose residue is not rate limiting. The contrast between the failure to glucosylate the unglucosylated peptide and the observation of glucosylation of those already glucosylated (Table 3) might be related to the question of whether the autoglycosylation is intra- or inter-molecular. Our results, and those previously reported for the *p*-nitrophenyl saccharides,[35] show that intermolecular transglucosylation to an already glucosylated acceptor substrate can occur. The kinetics of glucosylation when glycogenin glucosylates itself are compatible also with an intermolecular reaction.[38] However, it is possible that the addition of the very first glucose residue, to Tyr-194, is intramolecular, explaining the continued failure (Table 3) to observe intermolecular glucosylation of an unglucosylated peptide.

**Table 3** Acceptor activity in [$^{14}$C]glucosyl transfer by recombinant glycogenin

| Acceptor | % Acceptor undergoing glucosylation | |
|---|---|---|
| | 2 μmol $dm^{-3}$ UDP-glucose[a] | 150 μmol $dm^{-3}$ UDP-glucose[b] |
| α-Glucopeptide **5** | 0.0002 | 0.004 |
| β-Glucopeptide **6** | 0.0004 | 0.005 |
| α-Maltopeptide **7** | 0.003 | 0.010 |
| β-Maltopeptide **8** | 0.005 | 0.048 |
| Unglycosylated peptide **9** | 0.0000 | 0.000 |
| *p*-Nitrophenyl α-glucoside | 0.0007 | |
| *p*-Nitrophenyl α-maltoside | 0.0015 | |

[a] Concentration of peptide acceptors ~ 1 mmol $dm^{-3}$, and of *p*-nitrophenyl glycosides 10 mmol $dm^{-3}$. [b] Concentration of peptide acceptors ~ 0.5 mmol $dm^{-3}$.

The bond between the first glucose residue and Tyr-194 is believed to have the α-configuration. This follows from the bond's susceptibility to isoamylase, which splits the α-1,6-branch points of glycogen and also splits the bond between maltosaccharides and glycogenin.[17] Similarly, *p*-nitrophenyl α-maltoside is an acceptor in intermolecular glucosylation by glycogenin but the β-maltoside is not.[35] It was, therefore, unexpected when dodecyl β-maltoside proved to be a better acceptor than both *p*-nitrophenyl α-maltoside and dodecyl α-maltoside.[36] Similarly, in the present work, the peptides in which the glycosidic linkage to tyrosine is β are found to be superior as acceptors to those with an α linkage (Table 3). These contrasting results offer an intriguing, currently inexplicable, insight into the mechanism and specificity of glucose transfer by glycogenin.

## Conclusions

We have demonstrated the possibility of synthesizing tyrosine-glycosylated peptides by using glycosylated derivatives of

$N^{\alpha}$-Fmoc-Tyr-*O*Pfp as building blocks in solid-phase peptide synthesis. The synthetic glycopeptides were stable under the reaction conditions, including cleavage of the glycopeptides from the resin with 95% aq. TFA and deprotection of the carbohydrate moieties with sodium methoxide in methanol. By the use of real-time monitoring of the acylations it was shown that acylations towards the N-terminal proceeded slowly, indicating aggregation of the peptide on the resin. The rate of the $N^{\alpha}$-Fmoc cleavage was also slower in this part of the sequence. Both findings are in good agreement with the prediction of the formation of a β-strand in the N-terminal part of the peptides.

While none of the synthetic substrates were good acceptors for glucosyl transfer by recombinant glycogenin, the glycosylated peptides represent the first peptide substrates where successful intermolecular glucosylation has been achieved, and point to the possibility that self-glucosylation of already glucosylated glycogenin may also be intermolecular.[38] The failure to observe glucosylation of the corresponding glucose-free synthetic peptide might suggest that autoglycosylation of Tyr-194 in glycogenin is intramolecular.

## Experimental

HPLC-grade solvents were purchased from Labscan Ltd. (Dublin, Ireland). DMF was distilled by fractional vacuum distillation at 45 °C on a column of Raschig rings prior to use. Methanol was stored over 3 Å molecular sieves. Dhbt-OH and TBTU were purchased from Fluka (Switzerland), while the Rink-amide linker and suitably protected $N^{\alpha}$-Fmoc amino acids were purchased from NovaBiochem (Switzerland). Dhbt esters were prepared as previously described[39,40] or purchased from NovaBiochem. Methoxide solution (1 mol $dm^{-3}$) was prepared from sodium and dry methanol and diluted prior to use. Concentrations were performed under reduced pressure at < 35 °C (bath). Sequence analysis was performed on an ABI 470A Protein Sequenzer equipped with an ABI 120A PTH Analyser or on an ABI 477A Protein Sequenzer equipped with an ABI 120 Analyser. Amino acid composition was determined on a Pharmacia LKB Alpha Plus amino acid analyser, after hydrolysis with 6 mol $dm^{-3}$ HCl at 110 °C for 24 h. Preparative HPLC was carried out on a Waters system equipped with a Waters 991 photodiode array detector and a Delta Pak column (20 × 200 mm; 10 $cm^3$ $min^{-1}$) packed with reversed-phase $C_{18}$. Analytical chromatograms were obtained using a Waters RCM 8 × 10 module with a Waters 8 NV $C_{18}$ 4μ column (1 $cm^3$ $min^{-1}$). Buffer A was 0.1% aq. TFA, and buffer B was 0.1% TFA in 90% acetonitrile and 9.9% water. NMR spectra were recorded on a Bruker AMX-600 MHz spectrometer, and the experiments were carried out at 300 K. ESMS was performed in the positive mode on a VG Quattro mass spectrometer with 50% aq. acetonitrile as the liquid phase. Theoretical relative molecular masses are calculated as average masses. LSIMS was performed on a VG70SE instrument, using bis(hydroxyethyl) disulfide as matrix.

### Ac-[α-Glc-Tyr$^{194}$]Glycogenin(184–208)-$NH_2$ 5

Synthesis of compound **5** was performed in DMF with a custom-made, fully automated, continuous-flow peptide synthesizer (flow = 1.44 $cm^3$ $min^{-1}$) on PEGA-resin (0.20 g, 0.040 mmol). The resin was derivatized, on the column, with Rink-amide linker by coupling of the Fmoc derivative (54 mg, 0.10 mmol) using TBTU (30 mg, 0.095 mmol) in the presence of *N*-ethylmorpholine (NEM) (25 $mm^3$, 0.20 mmol). Before removal of the first Fmoc group, free amino groups were acetylated by circulation of acetic anhydride (50 $mm^3$, 12.5 equiv.) and DMAP (15 mg, 3 equiv.) for 20 min. The amino acids were coupled as $N^{\alpha}$-Fmoc-protected Dhbt esters (2.5 equiv., 0.10 mmol), except Asn, which was incorporated as $N^{\alpha}$-Fmoc-Asn(Trt)-OH (60 mg, 0.10 mmol) promoted by TBTU (30 mg, 0.095 mmol) in the presence of NEM (25 $mm^3$, 0.20 mmol). The hydroxy groups of Ser, Thr and unglycosylated Tyr were *tert*-butyl-protected, and the ε-amino group of Lys was Boc-protected. Fmoc groups were removed with 20% piperidine and the deprotection was monitored by measurement of the UV absorbance of the effluent at 312 nm. The α-glucosylated tyrosine building block **1**[21] (115 mg, 0.10 mmol) was coupled with addition of Dhbt-OH (6 mg, 0.04 mmol). Acylations were followed with a solid-phase spectrophotometer measuring the transmission through the column at 440 nm. However, owing to some persistent colour on the PEGA-resin, while monitoring fixed acylation times were used as follows: amino acid 1 (Lys) to 10 (Ala) 90 min, 11 (Pro) to 18 (Ile) 180 min, and 19 (Ser) to 25 (Phe) 360 min. The glycosylated tyrosine building block **1** was coupled for 18 h. Fmoc-deprotection times were as follows: Rink-amide linker and amino acid 1 (Lys) to 14 (Ser) 10 min, 15 (glyc-Tyr) to 23 (Tyr) 15 min, and 30 min for the last two amino acids. After coupling of the last amino acid (Phe), the amino terminal was acetylated for 90 min with Ac-*O*-Dhbt[33] (20 mg, 0.10 mmol). After completion of the synthesis the resin was washed with dichloromethane and dried on the lyophilizer. The glycopeptide was cleaved from the resin by treatment with 95% aq. TFA (40 $cm^3$) for 120 min, followed by filtration, and successive washing of the resin with 95% aq. TFA (4 × 40 $cm^3$) and with 95% aq. acetic acid (~200 $cm^3$, until the colour of the solution changed from yellow to off-white). The combined filtrates were then concentrated. The crude benzoylated peptide (83 mg, 61%) was suspended in dry methanol (25 $cm^3$) and 0.1 mol $dm^{-3}$ NaOMe (7.5 $cm^3$) was added until wetted pH-paper indicated pH 11. After 24 h the mixture was neutralized by addition of small pieces of solid $CO_2$. The crude deprotected glycopeptide was dissolved in a 1:1 mixture of acetic acid–water (10 $cm^3$), filtered, and purified by reversed-phase HPLC. This afforded the title compound **5** (5 mg, 7% from crude product). Amino acid analysis is presented in Table 1. $^1$H NMR data of the anomeric proton are presented in Table 2 [Found (ESMS): *m/z* 1514.6 $(M + H + K)^{2+}$, 1495.6 $(M + 2H)^{2+}$, 1017.7 $(M + H + Na + K)^{3+}$, 1010.3 $(M + 2H + K)^{3+}$, 1004.9 $(M + 2H + Na)^{3+}$, 767.7 $(M + 2H + 2K)^{4+}$ and 758.2 $(M + 3H + K)^{4+}$. $C_{143}H_{210}N_{30}O_{40}$ requires M, 2989.46].

### Ac-[β-Glc-Tyr$^{194}$]Glycogenin(184–208)-$NH_2$ 6

Compound **6** was synthesized with the glycosylated tyrosine building block **2**[21] (115 mg, 0.10 mmol) according to the same procedure as for compound **5** (see above). This gave the crude benzoylated glycopeptide (88 mg, 65%). Deprotection was performed by treatment with dry methanol (30 $cm^3$) and 0.1 mol $dm^{-3}$ NaOMe (10 $cm^3$) for 14 h. The solution was neutralized by addition of solid $CO_2$, and concentrated. The residue was not completely dissolved in a 1:1 mixture of acetic acid–water (20 $cm^3$) and had to be filtered before purification by HPLC. Purification afforded the title compound **6** (11 mg, 14% from crude product). Amino acid analysis is presented in Table 1. $^1$H NMR data of the anomeric proton are presented in Table 2 [Found (ESMS): *m/z* 1514.6 $(M + H + K)^{2+}$, 1495.7 $(M + 2H)^{2+}$, 1010.2 $(M + 2H + K)^{3+}$ and 1004.8 $(M + 2H + Na)^{3+}$].

### Ac-[α-Glc-(1→4)-α-Glc-Tyr$^{194}$]Glycogenin(184–208)-$NH_2$ 7

Synthesis of compound **7** was carried out on PEGA-resin (0.20 g, 0.040 mmol) using $N^{\alpha}$-Fmoc-protected Pfp esters of the amino acids in the coupling steps, with the addition of Dhbt-OH (6 mg, 0.040 mmol). Derivatization of the resin with Rink-amide linker, removal of the Fmoc-groups, protection of the side-chains of Ser, Thr, unglycosylated Tyr and Lys, acylation times and times for removal of the Fmoc groups were the same as for compound **5** (see above). In addition, the side-chain amide of Asn was Trt-protected. Before removal of the first Fmoc group, free amino groups were capped by circulation of

acetic anhydride (40 $mm^3$, 0.42 mmol) in DMF through the column for 10 min. The glycosylated tyrosine building block **3**[21] (65 mg, 0.040 mmol, 1 equiv.) was used, and it was coupled for 20 h with addition of Dhbt-OH (6 mg, 0.040 mmol). In case of incomplete coupling of building block **3**, acetylation was performed by circulation of a solution of acetic anhydride (40 $mm^3$, 0.42 mmol) in DMF for 25 min before Fmoc-deprotection. After Fmoc-deprotection of the last amino acid (Phe), the N-terminal was acetylated by circulation of acetic anhydride (200 $mm^3$, 2.1 mmol) for 65 min. When the synthesis was complete the resin was washed successively with DMF and dichloromethane and then dried on the lyophilizer. Deprotection of the side-chain protecting groups and cleavage of the glycopeptide from the resin was performed by treatment with 95% aq. TFA (20 $cm^3$) for 2 h. Filtration and washing successively with 95% aq. TFA (2 × 10 $cm^3$) and 95% aq. acetic acid (4 × 10 $cm^3$) was followed by concentration. The peptide was precipitated by addition of diethyl ether and after decantation the crude benzoylated product was isolated (110 mg, 71%). A sample of crude product (7 mg) was dissolved in 60% aq. TFA (1.5 $cm^3$) and purified by preparative HPLC (10 min 40% B, then a linear gradient of 40–60% B in 40 min). This gave protected (heptabenzoylated) title compound (4.4 mg, 63% from crude product) [Found (ESMS): *m/z* 1941.1 $(M + 2H)^{2+}$; 1306.9 $(M + 2H + K)^{3+}$ and 1294.4 $(M + 3H)^{3+}$. $C_{198}H_{248}N_{30}O_{52}$ requires M, 3880.32]. The truncated product, Ac-glycogenin(193–208)-$NH_2$, was also isolated (~0.1 mg, ~3% from crude product) [Found (ESMS): *m/z* 764.0 $(M + 2H)^{2+}$ and 783.0 $(M + H + K)^{2+}$. $C_{73}H_{108}N_{18}O_{18}$ requires M, 1525.8]. Debenzoylation of crude product (30 mg) was performed in methanol (30 $cm^3$) by treatment with 1 mol $dm^{-3}$ sodium methoxide in methanol (0.75 $cm^3$) for 48 h, followed by neutralization with solid $CO_2$ and concentration. The residue was dissolved in 50% aq. TFA (2.4 $cm^3$) and purified by preparative HPLC (10 min 30% B, then a linear gradient of 30–70% B in 80 min). This gave pure compound **7** (11 mg, 45% from crude product). Amino acid analysis is presented in Table 1. $^1H$ NMR data of the anomeric protons are presented in Table 2 [Found (ESMS): *m/z* 1587.5 $(M + H + Na)^{2+}$, 1576.4 $(M + 2H)^{2+}$, 1065.0 $(M + 2H + K)^{3+}$, 1058.8 $(M + 2H + Na)^{3+}$, 1072.8 $(M + 3Na)^{3+}$ and 798.7 $(M + 3H + K)^{4+}$. $C_{149}H_{220}N_{30}O_{45}$ requires M, 3151.56].

**Ac-[α-Glc-(1→4)-β-Glc-Tyr$^{194}$]Glycogenin(184–208)-$NH_2$ 8**

Compound **8** was synthesized according to the same procedure as for compound **5** (see above), using 2 equiv. of the glycosylated tyrosine building block **4**[21] (89 mg, 0.08 mmol). This afforded crude acetylated glycopeptide (75 mg, 54%). O-Deacetylation was performed in dry methanol (20 $cm^3$) and 1 mol $dm^{-3}$ NaOMe (1 $cm^3$) and after 21 h the solution was acidified by addition of 95% aq. acetic acid. The solvents were evaporated off and the residue was dissolved in 77% aq. acetic acid (7.5 $cm^3$) and purified by reversed-phase HPLC using 30% B for 20 min followed by a linear gradient from 30–50% B over a period of 80 min. This gave compound **8** (8 mg, 12% from crude product). Amino acid analysis is presented in Table 1. $^1H$ NMR data of the anomeric protons are presented in Table 2 [Found (ESMS): *m/z* 1598.7 $(M + 2Na)^{2+}$, 1576.3 $(M + 2H)^{2+}$, 1073.7 $(M + 2H + K)^{3+}$, 1058.9 $(M + 2H + Na)^{3+}$, 1073.7 $(M + 2H + K)^{3+}$ and 798.4 $(M + 3H + K)^{4+}$].

**Ac-Glycogenin(184–208)-$NH_2$ 9**

Synthesis of peptide **9** was performed on a Synergy Personal Peptide Synthesizer from Applied Biosystems using an Amide Fmoc Peptide Synthesis Column (0.025 mmol). The N-terminal was acetylated with Ac-*O*-Dhbt (15 mg, 0.075 mmol) on the synthesizer. The peptide was cleaved from the resin by treatment with 95% aq. TFA for 2 h, filtered off, and washed successively with 95% aq. TFA (5 × 1 $cm^3$) and 95% aq. acetic acid. Concentration, and precipitation with diethyl ether (15 $cm^3$) afforded the crude peptide (57 mg, 81%), which was dissolved in 60% aq. acetic acid and purified by HPLC using 20% B for 20 min followed by a linear gradient of 20–80% B for 120 min. This afforded the title compound **9** (8 mg, 14% from crude product). Amino acid analysis is presented in Table 1 [Found (ESMS): *m/z* 1436.1 $(M + 2Na)^{2+}$, 1425.1 $(M + H + Na)^{2+}$, 1414.4 $(M + 2H)^{2+}$, 958.1 $(M + H + 2Na)^{3+}$, 950.9 $(M + 2H + Na)^{3+}$ and 943.5 $(M + 3H)^{3+}$. $C_{137}H_{200}N_{30}O_{35}$ requires M, 2827.31].

**Assay for glucosyl transfer by recombinant glycogenin**

Recombinant rabbit-muscle glycogenin was prepared according to the method of Alonso *et al.*[15] In the first series of tests incubation mixtures contained 50 mmol $dm^{-3}$ Tris·HCl buffer, pH 7.4, 5 mmol $dm^{-3}$ $MnCl_2$, glycogenin (15 pmol), 130000 cpm of UDP[$^{14}$C]-glucose (196 pmol) and 390, 400, 330, 390 and 400 μg of compounds **5** (α-Glc), **6** (β-Glc), **7** [α-Glc-(1→4)-α-Glc], **8** [α-Glc-(1→4)-β-Glc] and **9** (unglycosylated), respectively, in a total volume of 0.1 $cm^3$. After incubation (3 h) at room temperature the samples were injected into Aquapor OD-300, 7μ cartridge for HPLC. The conditions for HPLC chromatography were as follows: solvent A, 0.05% TFA in water; solvent B, 0.05% TFA in acetonitrile, gradient 0–60% B during 60 min at a flow rate of 1 $cm^3$ $min^{-1}$. The elution profile was monitored by absorbency at 220 nm and the entire peak was collected as a single fraction, the volume of which was measured (2.0–2.5 $cm^3$). A 0.5 $cm^3$ portion was mixed with scintillation cocktail (7 $cm^3$) and counted in a liquid scintillation counter. *p*-Nitrophenyl α-glucoside and *p*-nitrophenyl α-maltoside (1 μmol of each) were tested in the same way. In the second series, the amount of UDP[$^{14}$C]-glucose was increased to 15 nmol (1.46 × $10^6$ cpm) and the tested acceptor substrates were **5** (190 μg), **6** (150 μg), **7** (151 μg), **8** (147 μg) and **9** (140 μg). The results are presented in Table 3.

## Acknowledgements

We thank Hanne Christiansen for synthesizing the unglycosylated peptide and Dr Ib Svendsen for the amino acid and sequence analyses. We also thank Professor William J. Whelan for his kind interest in this work and for valuable discussions. This work was supported by grants from the Danish Technical Research Council (K. J. J.) and from the US National Health Institutes of Health (DK 37500, J. L. and W. M. L.).

## References

1 J. Montreuil, *Adv. Carbohydr. Chem. Biochem.*, 1980, **37**, 157.
2 H. Lis and N. Sharon, *Eur. J. Biochem.*, 1993, **218**, 1.
3 S. Hase, H. Nishimura, S. Kawabata, S. Iwanaga and T. Ikenaka, *J. Biol. Chem.*, 1990, **265**, 1858.
4 K. Bock, J. Schuster-Kolbe, E. Altman, G. Allmaier, B. Stahl, R. Christian, U. B. Sleytr and P. Messner, *J. Biol. Chem.*, 1994, **269**, 7137.
5 J. Hofsteenge, D. R. Müller, T. Beer, A. Löffler, W. K. Richter and J. F. G. Vliegenthart, *Biochemistry*, 1994, **33**, 13 524.
6 P. S. Chen, H. K. Mitchell and M. Neuweg, *Insect Biochem.*, 1978, **8**, 279.
7 I. R. Rodriguez and W. J. Whelan, *Biochem. Biophys. Res. Commun.*, 1985, **132**, 829.
8 D. H. Williams, V. Rajananda, M. P. Williamson and G. Bojesen, *Top. Antibiot. Chem.*, 1980, **5**, 123.
9 K. J. Kramer, T. L. Hopkins, R. F. Ahmed, D. D. Mueller and G. Lookhart, *Arch. Biochem. Biophys.*, 1980, **205**, 146.
10 P.-W. Lu, K. J. Kramer, P. A. Seib, D. D. Mueller, R. F. Ahmed and T. L. Hopkins, *Insect Biochem.*, 1982, **12**, 377.
11 N. J. Skelton, M. M. Harding, R. J. Mortishire-Smith, S. K. Rahman, D. H. Williams, M. J. Rance and J. C. Ruddock, *J. Am. Chem. Soc.*, 1991, **113**, 7522.
12 P. Messner, R. Christian, J. Kolbe, G. Schulz and U. B. Sleytr, *J. Bacteriol.*, 1992, **174**, 2236.
13 M. Isobe, N. Kondo, K. Imai, O. Yamashita and T. Goto, *Agric. Biol. Chem.*, 1981, **45**, 687.
14 D. G. Campbell and P. Cohen, *Eur. J. Biochem.*, 1989, **185**, 119.

15 M. D. Alonso, J. Lomako, W. M. Lomako, W. J. Whelan and J. Preiss, *FEBS Lett.*, 1994, **352**, 222.
16 C. Smythe and P. Cohen, *Eur. J. Biochem.*, 1991, **200**, 625.
17 J. Lomako, W. M. Lomako and W. J. Whelan, *Carbohydr. Res.*, 1992, **227**, 331.
18 S. Horvat, L. Varga and J. Horvat, *Synthesis*, 1986, 209.
19 L. Varga, S. Horvat, C. Lemieux and P. W. Schiller, *Int. J. Pept. Protein Res.*, 1987, **30**, 371.
20 M. Meldal, in *Neoglycoconjugates: Preparation and Application*, ed. Y. C. Lee and R. T. Lee, Academic Press, San Diego, 1994, p. 149.
21 K. J. Jensen, M. Meldal and K. Bock, *J. Chem. Soc., Perkin Trans. 1*, 1993, 2119.
22 A. Vargas-Berenguel, M. Meldal, H. Paulsen, K. J. Jensen and K. Bock, *J. Chem. Soc., Perkin Trans. 1*, 1994, 3287.
23 S. Mouritsen, M. Meldal, J. Ruud-Hansen and O. Verdelin, *Scand. J. Immunol.*, 1991, **34**, 421.
24 M. Meldal, *Tetrahedron Lett.*, 1992, **33**, 3077.
25 F.-I. Auzanneau, M. Meldal and K. Bock, *J. Pept. Sci.*, 1995, **1**, 31.
26 H. Rink and B. Ernst, *Peptides 1990*, ed. E. Giralt and D. Andreu, ESCOM Science Publishers B.V., Leiden, 1991, p. 418.
27 R. Knorr, A. Trzeciak, W. Bannwarth and D. Gillessen, *Tetrahedron Lett.*, 1989, **30**, 1927.
28 L. Kisfaludy and I. Schön, *Synthesis*, 1983, 325.
29 L. R. Cameron, J. L. Holder, M. Meldal and R. C. Sheppard, *J. Chem. Soc., Perkin Trans. 1*, 1988, 2895.
30 L. Cameron, M. Meldal and R. C. Sheppard, *J. Chem. Soc., Chem. Commun.*, 1987, 270.
31 K. J. Jensen, Ph.D. Thesis, University of Copenhagen, 1992.
32 R. C. de L. Milton, S. C. F. Milton and P. A. Adams, *J. Am. Chem. Soc.*, 1990, **112**, 6039.
33 M. Meldal, T. Bielfeldt, S. Peters, K. J. Jensen, H. Paulsen and K. Bock, *Int. J. Pept. Protein Res.*, 1994, **43**, 529.
34 K. Bock and H. Thøgersen, *Annu. Rep. NMR Spectrosc.*, 1982, **13**, 1.
35 J. Lomako, W. M. Lomako and W. J. Whelan, *FEBS Lett.*, 1990, **264**, 13.
36 S. M. Manzella, L. Rodén and E. Meezan, *Glycobiology*, 1995, **5**, 263.
37 Y. Cao, A. M. Mahrenholz, A. A. DePaoli-Roach and P. J. Roach, *J. Biol. Chem.*, 1993, **268**, 14 687.
38 M. D. Alonso, J. Lomako, W. M. Lomako and W. J. Whelan, *J. Biol. Chem.*, 1995, **270**, 15 315.
39 E. Atherton, L. Cameron, M. Meldal and R. C. Sheppard, *J. Chem. Soc., Chem. Commun.*, 1986, 1763.
40 E. Atherton, J. L. Holder, M. Meldal, R. C. Sheppard and R. M. Valerio, *J. Chem. Soc., Perkin Trans. 1*, 1988, 2887.

*Paper* 5/06990D
*Received* 23*rd October* 1995
*Accepted* 24*th November* 1995